Consideration Related to the Aharonov–Bohm Effect

*Khang Jeong-woo, Khang Dong-shik**

Aharonov–Bohm Effect에 관한 고찰

康禎友, 康東植*

I. Introduction

In classical theory, forces need to be known only on the of the particles, therefore, classical theory is the local theory. These forces are expressed in terms of electromagnetic field only in electromagnetism (i.e. Lorentz force

$q[\vec{E}+\frac{1}{c}\vec{V}\times\vec{B}]$)

Thus, in all regions where field strengths are inaccessible, the motion of the particles is not affected by this Lorentz force, even though potentials exist which are considered as mathematical auxiliary quantities.

But in quantum theory, the situation is different from the above. Also, in quantum theory the physical quantities are all guage invariant. It may seem that potentials are not considered as basic quantities because they depend on a guage. But if potentials are expressed in a guage invariant form (i.e. $\oint\vec{A}\cdot d\vec{r}$), they are considered as basic quantities through guage invariant quantities(not through field strengths).

In connection with this, Aharonov–Bohm have pointed out some effects (particularly magnetic effect) of potentials in quantum theory through the double–slits experiments. In experiments Aharonov–Bohm predict an effect upon electrons passing outside a long solenoid. It is well known that outside an infinitely long solenoid the field strength is free and inside of it the field is parallel to the solenoid axis, but vector potentials exist in all regions.

Thus the fact that there is the Aharonov–Bohm effect in such a region where the field strength is free but vector potentials exist does

師範大學 助敎授, 師範大學 講師*

not imply that vector potentials play a role in this case. Of course, a role appears as an indirect gauge invariant quantity.

But in local quantum theory, interaction of electron with field strength happens at the point where the electron exists. Therefore only non–zero quantities correspond to the observable physical effect on the electron.

Since there is no way in quantum theory to describe the interaction of the electron with field strengths only in terms of field quantities, it seems that the potentials are basic quantity and the quantum theory is nonlocal one.

II. Review of the Aharonov–Bohm Effect

We are now in position to discuss the original form of the Aharonov–Bohm effect itself. Consider a particle of charge e passing above or below a very long impenetrable solenoid as shown in Figure I). In this arrangement the slits are a few wavelengths wide and only slightly divergent waves are used because the waves travel some part of distance between the slits.

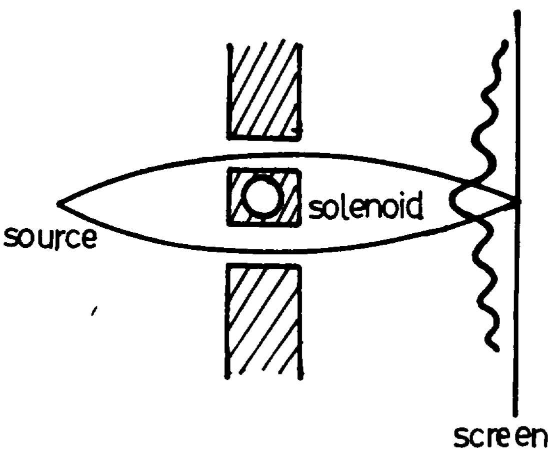

Figure I The Anaronov—Bohm interference pattern deflection

The electron de Broglie wave originating from the source is split into two parts and is coherently recombined to produce the interference pattern on the screen.

If the solenoid is not inserted into this arrangement, the interference pattern is exactly described by the classical theory except the fact that the electron de Broglie waves are at the origin in quantum theory.

How different is the situation, when the solenoid is involved? Of course, the regions where electron is passing above or below are inaccessible to the magnetic field; i.e. a field-free region. Therefore the electron de Broglie waves are not affected by Lorentz force, and according to the classical theory there is no change in the interference patteren on the screen.

A number of experiments, however, have showed that the interference pattern is unchanged but is shifted. This implies that in a region where the field strength is free the potential plays a role, producing the path difference and the interference pattern is shifted. Also, in this case, the potential is the vector potential.

The path difference is calculated quantitatively in terms of the WKB–method. The wave function of the electron de Broglie waves are expressed as

$$\Psi(\vec{x},t)=\sqrt{\rho}\ \exp\left[\frac{iS(\vec{x},t)}{\hbar}\right] \qquad (1)$$

where ρ is probability density and $\frac{S(\vec{x},t)}{\hbar}$ phase angle.

If one inserts this expression for the wave function of the electron de Broglie waves into the Schrödinger's wave equation, he gets the equation

$$\frac{1}{2m}|\nabla S(\vec{x},t)|^2\ V(\vec{x})+\frac{\partial S(\vec{x},t)}{\partial t}=0 \qquad (2)$$

This is known as Hamilton–Jacob equation in classical theory. where $S(\vec{x},t)$ stands for Hamilton's principal function. Thus one has a semi-classical interpretation of the phase of the wave function : $\hbar$ times the phase is equal to Hamilton's principal function provided that $\hbar$ can be considered as a small quantity. when the Hamiltonian is not a function of time, the Hamiltonian's principal function $S(\vec{x},t)$ can be written as follows.

$$S(\vec{x},t)=\bar{S}(\vec{x})-Et \tag{3}$$

where $\bar{S}(\vec{x})$ is time–independent function.

Inserting this equation into the equation(2) and restricting the problem to one dimension, the following solution of the Hamilton–Jacob equation is gotten.

$$\bar{S}(x) = \pm\int^{x} dx' \sqrt{2m[E-V(x')]} \tag{4}$$

If initially the electron enters a slit with the momentum $P=\sqrt{2mE}$ (i.e. $V(x)=0$) the potential. due to the solenoid. changes the momentum as follows.

$$P \longrightarrow P+\frac{e}{c}A_x \tag{5}$$

and the formula (4) is rewritten in terms of the changed momentum (5).

$$\bar{S}(x) = \pm\int^{x} dx' [\sqrt{2mE}+\frac{e}{c}A] \tag{6}$$

In 3–dimensional case this formulae is reexpressed

$$\bar{S}(\vec{x}) = \pm\int^{\vec{r}} [\vec{P}+\frac{e}{c}\vec{A}]\cdot d\vec{r}' \tag{7}$$

where the upper limit $\vec{r}$ stands for the path which electron de Broglie waves transverse. Thus the electron de Broglie waves passing through the different paths are described in terms of the wave functions

$$\begin{aligned}\Psi_{\vec{r}_1}(\vec{x},t)&=\sqrt{\rho}\,\exp\{i[\bar{S}_{\vec{r}_1}(\vec{x})-Et]/\hbar\}\\ \Psi_{\vec{r}_2}(\vec{x},t)&=\sqrt{\rho}\,\exp\{i[\bar{S}_{\vec{r}_2}(\vec{x})-Et]/\hbar\}\end{aligned} \tag{8}$$

respectively where subscript $\vec{r}_1, \vec{r}_2$ stands for the path and time t is same for the two different paths.

Therefore. the path difference due to the potential is

$$[S_{\vec{r}_2}(\vec{x})-S_{\vec{r}_1}(\vec{x})]/\hbar=\pm\frac{e}{\hbar c}[\int^{\vec{r}_2}\vec{A}\cdot d\vec{r}-\int^{\vec{r}_1}\vec{A}\cdot d\vec{r}]=\pm\frac{e}{\hbar c}\oint_c \vec{A}\cdot d\vec{r} \tag{9}$$

where integration path is the closed curve which encloses the solenoid.

This is just the result Ahsronov–Bohm predicts! Even if the potental $\vec{A}$ is changed by guage transformation integral $\oint\vec{A}\cdot d\vec{r}$ is guage invariant.

Ⅲ. Gauge Invariance and Locality of the Aharonov—Bohm Effect.

The classical electromangnetism is described by the four Maxwell's equation written in a differential form.

$$\begin{aligned}&\nabla\cdot\vec{D}=4\pi\rho \qquad \nabla\times\vec{H}=\frac{4\pi}{c}\vec{J}+\frac{1}{c}\frac{\partial\vec{D}}{\partial t}\\ &\nabla\cdot\vec{B}=0 \qquad \nabla\times\vec{E}+\frac{1}{c}\frac{\partial\vec{B}}{\partial t}=0\end{aligned} \tag{10}$$

When these equations are combined with the Lorentz force equation and Newton's second law of motion they completely described the classical dynamics of interacting charged particles and electromagnetic fields. As scalar potential Φ and vector potential $\vec{A}$ are introduced, the Maxwell equations (10) are rewritten as a smaller number of second–order partial differential equations.

$$\begin{aligned}&\nabla^2\Phi+\frac{1}{c}\frac{1}{\partial t}(\nabla\cdot\vec{A})=-4\pi\rho\\ &\nabla^2\vec{A}-\frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2}-\nabla(\nabla\cdot\vec{A}+\frac{1}{c}\frac{\partial\Phi}{\partial t})=\frac{4\pi}{c}\vec{J}\end{aligned} \tag{11}$$

Though the four Maxwell equations are now reduced to the set of two potential equations, they are still coupled equations. The magnetic induction is written in terms of the curl of vector potential, therefore even though the gradient of some scalar function is added the magnetic induction is left unchanged.
That is, under the transformation

$$\vec{A} \to \vec{A}' = \vec{A} + \nabla\Lambda \tag{12}$$

field strength is unchanged.

Thus chosing the appropriate scalar function Λ, one gets the following condition, i.e. radiation guage.

$$\nabla \cdot \vec{A} = 0. \tag{13}$$

Then the set of two potential equations is rewritten in the use of this radiation gauge.

$$\begin{aligned} &\nabla^2 \Phi = -4\pi\rho \\ &\nabla^2 \vec{A} - \frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} = -\frac{4\pi}{c}\vec{J} + \frac{1}{c}\nabla\frac{\partial \Phi}{\partial t}. \end{aligned} \tag{14}$$

It is well known that the electromagnetic waves propagate with a finite velocity. But the equations (14) imply that the scalar potential propagates instantaneously everywhere in space, and the vector potential propagates with a finite velocity C. It is, however, not the potentials that concern us.

Thus in classical theory the field strengths are regarded as basic quantities and the potentials as mathematical auxiliary quantities. In quantum theory, this situation is changed. The Lorentz force in classical theory is replaced by the average force in semiclassical quantum theory through the correspondence principle, i.e.

$$\langle \vec{F} \rangle = \langle \Psi | [-\nabla V + e\nabla\Phi - \frac{e}{c}\vec{V} \times \vec{B}] | \Psi \rangle \tag{15}$$

where V is the non-electrostatic potential, Φ scalar potential, $\vec{B} = \nabla \times \vec{A}$ magnetic induction, $\vec{V}$ the velocity operator of the matter considered, Ψ the wave function of that matter.

There is a difference between formula (15) and the classical Lorentz force $(\vec{F} = -\nabla V + e\nabla\Phi - \frac{e}{c}\vec{V} \times \vec{B})$ The formula (15) involves wave function Ψ and this wave function Ψ is obtained by solving the Schrödinger's wave equations, in turn the Schrödinger's equations have a following Hamiltonian.

$$H = (\vec{P} - \frac{e}{c}\vec{A})^2 + V. \tag{16}$$

Borrowing the Aharonov–Bohm's statements, "because there is no way in quantum mechanics to express the interaction of the matter with the electromagnetic field in terms of field equations the wave function entering into equation for the average force therefore cannot, in general, be known unless one first knows the potentials."

The formula (16) corresponds to the transformation

$$\vec{P} \to \vec{P} - \frac{e\vec{A}}{c} \tag{17}$$

All observable quantities in quantum theory are invariant under the guage transformation. Hamiltonian (16), therefore, must be guage invariant.

Then, under the gauge transformation (12) how does the wave function transform?

First, we construct an unitary representation U

$$\Psi \to \Psi' = U\Psi \tag{18}$$

and from the fact that Hamiltonian (16) is invariant under the guage transformation (12),

$$U^{+}(\vec{P} - \frac{e}{c}\vec{A} - \frac{e\nabla\Lambda}{c})U = \vec{P} - \frac{e\vec{A}}{c} \tag{19}$$

the unitary representation U which satisfies this condition (19) is

$$U = \exp\left[\frac{ie\Lambda}{\hbar c}\right] \tag{20}$$

and the transformed wave function is written

$$\Psi' = \exp\left[\frac{ie\Lambda}{\hbar c}\right]\Psi. \tag{21}$$

The untransformed wave function Ψ is, therefore, not gauge invariant. Using the parallel displacement,

$$\left[1-\frac{ie}{\hbar c}\vec{A}\cdot d\vec{r}+O(d\vec{r}^{2})\right]\Big|_{\text{at } \vec{r}+d\vec{r}} \tag{22}$$

the wave function which is gauge invariant is constructed.

This procedure is obtained as follows. Cosider some function of position $\vec{r}:\Phi(\vec{r})$. When the vector potential is free,

$$\Phi(\vec{r}+d\vec{r})\simeq\Phi(\vec{r})+(\nabla\Phi)\cdot d\vec{r}. \tag{23}$$

If the vector potential $\vec{A}$ is turned on.

$$\nabla\to\nabla-\frac{ie}{\hbar c}\vec{A}$$

$$\begin{aligned}\Phi'(\vec{r}+d\vec{r})&\simeq\Phi(\vec{r})+\left(\nabla-\frac{ie}{\hbar c}\vec{A}\right)\Phi\cdot d\vec{r}\\&\simeq\Phi(\vec{r}+d\vec{r})-\frac{ie}{\hbar c}\Phi\vec{A}\cdot d\vec{r}.\end{aligned} \tag{24}$$

Thus we obtain

$$\Phi' = \exp\left[-\frac{ie}{\hbar c}\vec{A}\cdot d\vec{r}\right]\Phi. \tag{25}$$

The guage invariant wave function is then written by

$$\Psi = \exp\left[-\frac{ie}{\hbar c}\vec{A}\cdot d\vec{r}\right]\Psi. \tag{26}$$

For a finite displacement along the path C,

$$\Psi = \exp\left[-\frac{ie}{\hbar c}\int_C\vec{A}\cdot d\vec{r}\right]\Psi \tag{27}$$

It also be used in describing the electron de Broglie wave because it differs only in phase from the original wave function.

The problem arises at this point. If the quantum theory is local one, then this formulae (27) is used in a naive form and as a result, the potentials (in this case, vector potential) are regarded as basic quantities.

Alternatively if quantum theory is non–local, the vector potential must be reexpressed in terms of magnetic induction $\vec{B}$ through

$$\nabla\times\vec{B}=\nabla(\nabla\cdot\vec{A})-\nabla^{2}\vec{A}. \tag{28}$$

Using the radiation gauge, the vector potential is

$$\vec{A}(\vec{r}_1)=\frac{1}{4\pi}\int\frac{(\nabla r_2\times\vec{B})\cdot d\vec{r}_2}{|\vec{r}_1-\vec{r}_2|}. \tag{29}$$

Since at the right hand side of the equation (29) the magnetic induction (i.e. field strength) only appears, the interaction of the electrons with field strength must be described non–locally and the situation becomes, therefore, more difficult.

For a while, let us discuss the local theory. In a local theory there is an assumption such as idealization of space–time measurement in arbitrarily small regions, and the field functions are continuous functions having the continuous parameter.

Therefore in a microscopic world it may seem that a local theory is re–constructed in a different way. There is, however, no concrete evidence of a discontinuity at small distances. Up to now there is no violation of the above assumption down to distance $\hbar c/\sqrt{S}\sim 10^{-15}$cm (where S is the squared center–of–mass energy). Also as required by Lorentz invariance and a microcausality, the velocity of the disturbance must not exceed the velocity of light, and the field functions relative to spacelike separated intervals commute, that is,

$$[\Phi(x), \Phi(y)] = 0 \quad \text{for} \quad (x-y)^2 < 0 \qquad (30)$$

where parameter x, y is four vector and we used the metric (1, −1, −1, −1). Another reason which one needs the local theory is that there exists no different satisfactory theory.

Even though there exists an alternative theory, it must be approximated to the local theory in large distance regions.

Thus one can use the local theory down to ~ 10^{-15}cm distances and (in this regions quantum theory is local one.) one can conclude that the potentials are basic quantities. Of course, even though this is true in quantum theory, in classical theory the field strengths are still basisc quantities.

Ⅳ. Conclusion

In contrast with the classical theory, in which field strengths play a fundamental role in deseribing the variety of nature, the potentials are basic quantities in quantum theory.

The Aharonov–Bohm effect supports this fact strongly. In quantum theory Hamiltonian, when the potential exists, is expressed the formula (16) and, this potential (in this case vector potential) acts as an operator and changes the eigenvalue of the Hamiltonian. Because we do not know how to express the interaction of the matter with electromagenetism solely in terms of the field strengths. as seen the above discussion we conclude that the potentials play a basic role in quantum theory.

Of course the potential can be expressed in terms of the field strength. In this case theory becomes the nonlocal theory. Though there are some evidence supporting the nonlocal theory (i.e. Bell's inequality), the local theory is still the strong method describing the physical phenomena.

Therefore, as for as the local theory is core we must regard the potential as the basic quantity in quantum theory.

Literature cited

Y. Aharonov and D. Bohm. Significance of Electromagnetic potentials in the Quantum theory, Phys. Rev. 115, 485 (1959), Phys. Rev. 123, 1511(1951), Pys. Rev. 125, 2192(1962).

W. H. Furry and N. F. Ramsey, Significance of Potentials in quantum Theory, Phys. Rev. 118, 623 (1960).

Bryce S. DeWitt, Quantum Theory without Electromagnetic Potentials, Phys. Rev. 15, 2189 (1962).

U. Klein, Comment on Condition for nonexistence of Aharonov–Bohm effect. Phys. Rev. 23D, 1463(1981).

T. H. Voyer, Classical Electromagnetic Deflections and lag Effect Associated with Quautum Interference Pattern Shifts: Considerations Related to the Aharonov–Bohm Effect. Phys. Rev. 8D, 1667(1973), Phys. Rev. 8D, 16879 (1973).

D. M. Greenberger, Reality and significance of the Aharonov–Bohm Effect, Phys. Rev. 23d, 1460(1981).

J. G. Cramer, The Transactional Interpretation of Quantum Mechanics, Rev. Mod. Phys. 58, 647(1986)

T. T. wu and C. N. Yang, Concepts of Nonintegrable Phase Factors and Global Formulation of Gauge Fields, Phys. Rev. 12d, 3845 (1975).

J. D. Bjorken and S. D. Drell, "Relativistic Quantum Fields" 1965, McGraw–Hill, Inc., p. 3.

J. J. Sakurai, Modern Quantum Mechanics, 1985, The Benjamin/Curmmings Publishing Company, Inc. p. 97.

P. Ramond, Field Theory A Modern Primer, 1981, The Benjamin/Cummings Publishing Company, Inc. p.276.

C. Itzykson and J. B. Zuber Quantum Field Theory, 1980, McGraw–Hill, Inc., p. 105.

H. Goldstein, Classical Mechanics, 2nd Edition, Addison–Wesley Publishing Company, p.438.

國 文 抄 錄

Aharonov–Bohm 효과를 WKB 방법을 사용해서 설명하고, 국소장 이론으로 자연을 서술하는 방법의 타당성에 대해 고찰하였다.